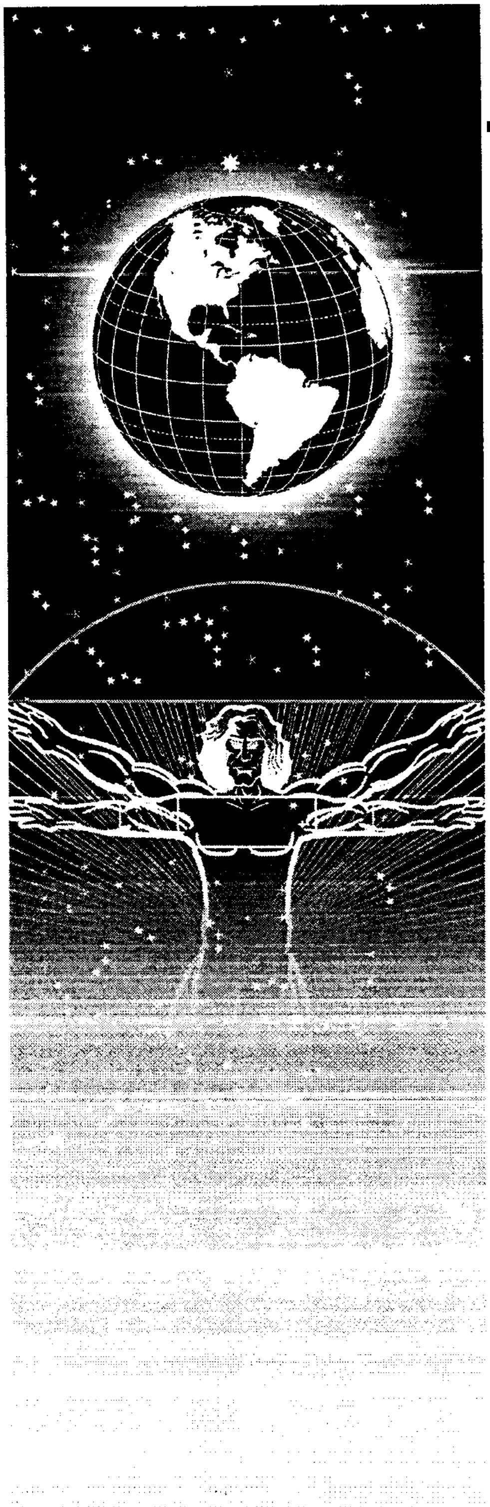
AL/HR-TP-1996-0037



UNITED STATES AIR FORCE
ARMSTRONG LABORATORY

Training Benefits Of Interactive Air Combat Simuation

Wayne L. Waag

Herbert H. Bell

March 1997

Approved for public release; distribution is unlimited.

Human Resources Directorate
Aircrew Training Research Division
6001 South Power Road, Bldg 558
Mesa Arizona 85206-0904

NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Office of Public Affairs has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This report has been reviewed and is approved for publication.

HERBERT H. BELL
Project Scientist

ELIZABETH L. MARTIN
Technical Director

LYNN A. CARROLL, Colonel, USAF
Chief, Aircrew Training Research Division

Please notify AL/HRPP, 7909 Lindbergh Drive, Brooks AFB, TX 78235-5352, if your address changes, or if you no longer want to receive our technical reports. You may write or call the STINFO Office at DSN 240-3853 or commercial (210) 536-3853.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 1997	Final Report - January 1994 to November 1994

4. TITLE AND SUBTITLE

Training Benefits of Interactive Air Combat Simulation

5. FUNDING NUMBERS

PE - 62202F
PR - 1123
TA - B3
WU - 02

6. AUTHOR(S)

Wayne L. Waag
Herbert H. Bell

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Armstrong Laboratory
Human Resources Directorate
Aircrew Training Research Division
6001 South Power Road, Building 558
Mesa, AZ 85206-0904

8. PERFORMING ORGANIZATION REPORT NUMBER

AL/HR-TR-1996-0037

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Armstrong Laboratory
Human Resources Directorate
7909 Lindbergh Drive
Brooks Air Force Base, Texas 78235-5352

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

Armstrong Laboratory Technical Monitor: Dr Herbert H. Bell, (602) 988-6561.
This paper documents an invited paper which was presented at the Royal Aeronautical Society Flight Simulation Symposium held 16-17 November 1994 in London, England.

12a. DISTRIBUTION/AVAILABILITY STATEMENT	12b. DISTRIBUTION CODE
Approved for public release; distribution is unlimited.	

13. ABSTRACT *(Maximum 200 words)*

This paper presents the results of an investigation of the training benefits of distributed, interactive air combat simulation. The Multiship Research and Development (MULTIRAD) simulation facility located at Armstrong Laboratory's Aircrew Training Research Division in Mesa, Arizona was used to conduct this investigation. It permitted two F-15s to fly against a suite of manned and unmanned adversaries in a realistic combat environment. Simulation components represent independent subsystems operating as a part of a secure distributed simulation network. This local area network was connected to an air weapons controller simulator located at Brooks Air Force Base, Texas, using a dedicated T-1 telephone line. A week-long syllabus was designed consisting of nine sorties with four engagements per sortie. A building-block approach was taken so that scenarios increased in difficulty over the week. Sixty-three mission-ready (MR) F-15 pilots participated in the research effort. Critical incident/event data and ratings of situation awareness (SA) were gathered using two trained observers. Using the ratings from the two trainer observers, performance was found to improve for identical engagements flown early and late in the syllabus. At the end of the week, questionnaires were administered which gathered data on the perceived value of the simulation for training and how it might best be employed. Overall, very positive opinions were expressed by the research participants regarding the potential value of multiship simulation for training SA skills. These findings, while not providing definitive evidence of transfer of training to the air, nonetheless, do suggest that multiship simulation may be an effective tool for training SA skills for fighter operations.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Air Combat; Air combat simulation; Distributed interactive air combat simulation; Flight simulation; Flight training; MULTIRAD; Multiship simulation; SA; Simulation; Situation awareness; Training; Training Evaluation	17
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION ABSTRACT
Unclassified	Unclassified	Unclassified	UL

PREFACE

This effort documents work conducted at the Multiship Research and Development (MULTIRAD) simulation facility located at Armstrong Laboratory's Aircrew Training Research Division (AL/HRA) in Mesa, AZ. This paper documents an invited paper which was presented to the Royal Aeronautical Society Flight Simulation Symposium which was held in London, England from 16-17 November 1994. It discusses results on an investigation of the training benefits of distributed, interactive air combat simulation.

This effort was conducted under Work Unit 1123-B3-02, Tools for Assessing Situational Awareness. The principal investigator for this work unit was Dr Wayne L. Waag (retired). The current principal investigator is Dr Herbert H. Bell.

THE TRAINING BENEFITS OF INTERACTIVE AIR COMBAT SIMULATION

Wayne L. Waag, Ph.D. and Herbert H. Bell, Ph.D.

Aircrew Training Research Division
Armstrong Laboratory
Mesa Arizona, USA

Abstract

This paper presents the results of a recent investigation of the training benefits of distributed, interactive air combat simulation. The Aircrew Training Research Division's Multiship Research and Development (MULTIRAD) simulation facility, located at Mesa, AZ, was used which permitted two F-15s to fly against a suite of manned and unmanned adversaries in a realistic combat environment. Simulation components represent independent subsystems operating as part of a secure distributed simulation network. This local area network was connected to an air weapons controller simulator located at Brooks Air Force Base, Texas using a dedicated T-1 telephone line. A week-long syllabus was designed consisting of 9 sorties with 4 engagements per sortie. A building block approach was taken so that scenarios increased in difficulty over the week. Sixty-three mission-ready (MR) F-15 pilots participated in the study. Critical incident/event data and ratings of situation awareness (SA) were gathered using two trained observers. Using the ratings from the two trained observers, performance was found to improve for identical engagements flown early and late in the syllabus. At the end of the week, questionnaires were administered which gathered data on the perceived value of the simulation for training and how it might best be employed. Overall, very positive opinions were expressed by the study participants regarding the potential value of multiship simulation for training SA skills. These findings, while not providing definitive evidence of transfer of training to the air, nonetheless, do suggest that multiship simulation may be an effective tool for training SA skills for fighter operations.

Introduction

The United States Air Force (USAF) spends a great deal of money to develop and maintain the combat proficiency of its pilots. Most of this combat-oriented training is conducted at the operational unit as part of it's continuation training program. The basic instructional media for continuation training are the aircraft, the environment in which it operates, and the mission debrief. Together they provide an on-the-job training environment built around the opportunities for in-flight training. In-flight training opportunities, however, are limited by many factors (1). These factors include: peacetime training rules, resource limitation, technical constraints, and security restrictions. Each of these factors places restrictions or imposes unnatural constraints on training. Peacetime training rules impose altitude and weather restrictions, limit use of communications jamming, and require a minimum separation between aircraft. Resource limitations restrict the number of aircraft available for training, the number of flying hours available, and the size of the training ranges. Technical constraints limit the use of electronic warfare systems, prevent practice against an integrated air defense system, and limit the measurement of combat performance. Security restrictions prevent full employment of classified systems, communications, and tactics. These factors combine to limit the opportunities for training combat tasks at both individual and team levels.

Over the past four years, the Armstrong Laboratory's Aircrew Training Research Division, in cooperation with the Air Combat Command, has surveyed over 300 MR pilots and air weapons controllers (AWCs) to identify continuation training needs (2,3). Responses to these surveys were surprisingly similar no matter the respondent's experience level, unit, or weapon system. The consensus is that it is difficult to train the pilot and AWC to make full use of the weapon system as part of a combat team. Table 1 shows the combat training areas most frequently mentioned as needing improvement.

These mission areas involve the very tasks for which in-flight training is most likely to be constrained by the factors mentioned above. If anything, the negative impacts of these factors on training will increase in the future. Therefore, we must develop other training approaches that will maintain the readiness of our combat air forces. Simulation is one such approach (4). In particular, distributed interactive simulation seems especially promising since it offers the potential interactivity that characterizes the combat environment.

Table 1. Mission Activities Most Frequently Mentioned As Requiring Additional Training

Multibogey, four or more
All-aspect defense
Reaction to surface-to-air missiles
Dissimilar air combat tactics
Four-ship tactics
Reaction to air interceptors
Employment of electronic countermeasures
Chaff/flares employment

Because of the high cost of flight simulators and the potential consequences of inadequate training, one would assume there is an extensive research base establishing the value of training combat tasks in simulators. It is not unreasonable to ask questions such as: Was the simulator training effective? Can it be improved? How frequently is it needed? Is simulation worth the costs? All of these questions reflect the need to evaluate the potential benefits of distributed simulation for combat-oriented training.

Measuring Training Benefits

An immediate question becomes exactly how to evaluate the benefits of simulation training in preparation for a combat environment. Bell and Waag (5) have proposed a five stage sequential model which is briefly summarized.

Stage 1. Utility Evaluation. The objectives of the initial stage are to (a) evaluate the accuracy or fidelity of the simulation environment; and (b) to gather opinions concerning the potential value of the simulation within a training environment. These objectives are quite similar to those of operational test and evaluations (OT&Es) that are routinely conducted for most simulator acquisitions.

Stage 2. Performance Improvement. The objective of the second stage of the evaluation is to determine the extent to which performance improves during the course of training within the simulation environment. The major challenge during this stage of the evaluation is to ensure that there is a proper means of establishing that performance has indeed improved as a result of the training. This requires the development of mission scenarios that are flown before and after the training that are similar to but not identical to missions flown during training. It also requires the development and use of measures whereby improvements in performance can be meaningfully reflected.

Stage 3. Transfer to Alternative Simulation Environment. The question of generalizability now is raised--does training transfer to another environment? While the acid test is usually considered to be transfer to the air, it is our view that it is necessary to demonstrate transfer to other simulation environments as well. Recall that one of the primary justifications for multiplayer air combat simulation is the ability to practice certain events under conditions that are generally not available in the real world--short of war. Because of safety restrictions, security considerations, rules of engagement, etc., peacetime exercises will always be limited in terms of their situational fidelity. For this reason, it is essential that transfer be demonstrated to another simulation environment in which a wartime environment can be created.

Stage 4. Transfer to Flight Environment. If positive transfer to a wartime environment using another simulation facility has been shown, the next stage is to show transfer to the air. To some extent, such a transfer test is limited by the large number of peacetime restrictions that characterize current flight operations. For this reason, a smaller sample of the air environment would be selected for evaluation of transfer. To whatever extent possible, the transfer test should represent a highly controlled flight environment wherein performance data can be gathered easily.

Stage 5. Extrapolation to Combat Environment. The last stage of the evaluation process attempts to answer the question of the military value of the simulation training. As might be expected, an empirical approach is not amenable for this question. Rather, a modeling approach is recommended as a potential vehicle for extrapolating the potential value of the simulation training to a combat environment. An example of such an approach is provided by Deitchman (6) in an attempt to project the impact of training into a central European type of wartime scenario. In that case, arbitrary estimates were used to represent the potential impacts of training. For example, one might assume that target identification rate could be doubled through training. Using a simulation model, the impact of such changes on model behavior could be determined.

Previous Evaluations

In concert with this model, the Armstrong Laboratory has been gathering data over the past few years attempting to establish the value of simulation for air combat training. In 1988, a program was initiated with the Tactical Air Command (now Air Combat Command) to evaluate multiship air combat training using commercially available contractor facilities. In all, three studies were conducted. In the first study (7), 42 mission-ready F-15 pilots and 16 AWCs received four days of training at the McDonnell Douglas Aircraft simulation facility in St. Louis, Missouri. The training unit was the team comprised of two pilots (lead/wingman) plus the AWC. This team flew a variety of combat missions against an opposing force comprised of four to eight adversaries plus the adversary AWC.

Upon completion of training, pilots rated the value of both their "unit training" and the "simulation training" for 41 air-to-air tasks. The pilots felt that simulator training was much better than their current unit training for many air combat tasks including multibogey, chaff and flares employment, all-aspect defense, use of electronic countermeasures and counter-countermeasures, communications jamming, and work with the AWC. These tasks were also rated high in "need for additional training" prior to the start of simulator training. On the other hand, tasks such as air combat maneuvering (ACM), visual lookout, gun employment, and basic fighter maneuvering (BFM) were rated as better trained in their in-flight continuation training program than in the simulation. Air weapons controllers, however, rated all tasks as better trained in the simulation environment. Open-ended opinion data were also gathered, the results being quite positive toward the training.

In the second study, a follow-on evaluation was conducted using the same procedure but with a larger and more representative sample of pilots and AWCs (3). This evaluation produced essentially the same results. Based on the high user acceptance demonstrated during these utility evaluations, Air Combat Command continued this program under its own sponsorship.

In the third study, again using the same facility, in-simulator learning was also shown, in addition to positive user opinion. Subjects consisted of 16 teams, each team being made up of two pilots and an AWC. Each of the elements flew controlled offensive and defensive scenarios "before" and "after" three days of intensive simulation training. Digital data as well as videotapes of displays used for replay and debriefing purposes were archived and are currently being analyzed. Analyses revealed the teams' post-training mission effectiveness and survivability to be significantly higher than their pretraining scores.

Background to Current Effort

The impetus for the present investigation came directly from the US Air Force Chief of Staff. In 1991, he posed a series of questions concerning SA within the operational F-15 fighter world. First of all, What is SA? Can it be objectively measured? Is SA learned or does it represent a basic ability or characteristic that some pilots have and others do not? In response to the question, "what is it?" a working group at the Air Staff produced the following operator's definition of SA: "a pilot's continuous perception of self and aircraft in relation to the dynamic environment of flight, threats, and mission, and the ability to forecast, then execute tasks based on that perception (8)." While other definitions of SA within the literature focus primarily on processes underlying the assessment and resulting knowledge of the situation (9,10), our working definition also included forecasting, decision making, and task execution. From an operational Air Force perspective, SA is more than simply knowledge and understanding of the environment.

The Armstrong Laboratory subsequently initiated a research investigation that had three goals: first, to develop and validate tools for reliably measuring SA; second, to identify basic cognitive and psychomotor abilities that are associated with pilots judged to have good SA; and third, to determine if SA can be learned, and if so, to identify areas where cost-effective training tools might be developed and employed.

To develop measurement tools (the first goal of the study), it was first necessary to identify and describe critical behavioral indicators of the fighter pilot's ability to maintain good SA and successfully complete his mission. To this end, a cognitive task analysis of a typical F-15 air combat mission was conducted (11). The resulting analysis identified the significant types of decisions required of the flight members, the information required for making these decisions, and the observable activities the flight members performed to acquire this information. The results were further analyzed by an experienced fighter pilot to identify behavioral indicators considered most essential to SA. This subject matter expert (SME) emphasized that these behavioral indicators must be observable in the context of day-to-day squadron training activities and subject to evaluation by fighter pilots both in terms of their own performance and that of others. As a result of this analysis, 24 behavioral indicators organized in seven categories were identified and are shown in Table 2.

Based principally upon these behavioral indicators, a number of SA Rating Scales (SARS) were developed to measure SA in operational units. They were administered to 238 mission-ready F-15 pilots from 11 operational squadrons. From the SARS, a composite measure of SA was derived and found to be highly related to previous flight experience and current flight qualification (12).

Table 2. Behavioral Indicators and Categories of Performance

1. TACTICAL GAME PLAN
   - Developing plan
   - Executing plan
   - Adjusting plan on-the-fly
2. SYSTEM OPERATION
   - Radar
   - Tactical electronic warfare system
   - Overall weapons system proficiency
3. COMMUNICATION
   - Quality (brevity, accuracy, timeliness)
   - Ability to effectively use information
4. INFORMATION INTERPRETATION
   - Interpreting vertical situation display
   - Interpreting threat warning system
   - Ability to use controller information
   - Integrating overall information
   - Radar sorting
   - Analyzing engagement geometry
   - Threat prioritization
5. TACTICAL EMPLOYMENT-BVR
   - Targeting decisions
   - Fire-point selection
6. TACTICAL EMPLOYMENT-VISUAL
   - Maintain track of bogeys/friendlies
   - Threat evaluation
   - Weapons employment
7. TACTICAL EMPLOYMENT-GENERAL
   - Assessing offensiveness/defensiveness
   - Lookout
   - Defensive reaction
   - Mutual Support

These measures were used for two purposes. First, they served as a criterion measure against which to validate a battery of basic ability tests considered relevant to SA, thereby addressing the question of basic human abilities (the second goal of the study). The Situation Awareness Assessment Battery (SAAB), consisting of 24 computer-based tests of basic cognitive and psychomotor abilities (13), was also administered to the same sample of pilots at their home units.

Second, these measures served as a means of selecting a sample of pilots who participated in the simulation phase of the effort, in which performance was observed under realistic combat conditions. During this phase, simulated air combat mission scenarios were developed for assessing SA and a variety of performance measures gathered in an attempt to determine whether SA could be measured in simulation environment. Moreover, an attempt was made to examine the potential of this type of simulation for training critical SA skills. This paper presents some preliminary findings of the data gathered from a simulated air combat environment.

Method

Subjects

A total of 40 MR F-15 pilots, who were flight lead-qualified served as subjects. An additional 23 MR F-15 pilots served as wingmen throughout the data collection which began in Mar 93 and was completed in Jan 94.

Simulation System

The Armstrong Laboratory multiship simulation facility (MULTIRAD) located at Williams Air Force Base (WAFB), Arizona (now Williams Gateway Airport, Mesa, AZ) was used. The major components of the simulation system are shown in Figure 1. These components represent independent subsystems operating as part of a secure distributed simulation network. This local area network was connected to the air weapons controller simulator (AESOP) at Brooks Air Force Base (BAFB), TX by a dedicated T-1 telephone line. Additional details concerning the basic simulation architecture and components are available (14,15).

The manned flight simulators consisted of two F-15C simulators and two F-16 simulators. The F-15C simulators had high fidelity aerodynamic, engine, avionics, radio, sensor, and weapons simulations. Each F-15C simulator was equipped with an out-the-window visual display system covering approximately 360 deg horizontal by 200 deg vertical. The external visual scene was created using computer-generated imagery. The manned F-16 simulators had less fidelity and played the role of enemy aircraft in conjunction with computer-controlled adversaries. The visual and electronic signatures of these F-16 simulators were modified so that they appeared as the appropriate threat aircraft. Each F-16 simulator was equipped with a single channel of out-the-window visual imagery covering approximately 45 deg horizontal by 45 deg vertical.

A manned AWC provided the F-15C pilots with appropriate threat information and warnings. Depending upon the availability of qualified AWCs and equipment status, the AWC was either located at WAFB or BAFB. In either case, the AWC had a realistic simulation of the appropriate AWC console and communicated with the F-15C pilots by radio.

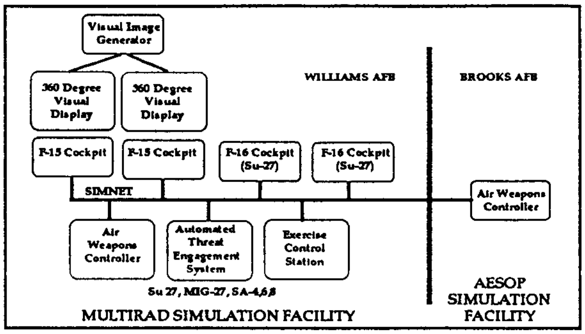


Figure 1. Multiship Simulation Facility

The exercise control system (ECS) consisted of a central console with the hardware and software necessary to create, start, observe, record, and stop the simulated air combat sorties. The SMEs who served as test directors and observers viewed monitors that provided a real-time view of each sortie. These monitors provided: 1) a plan view display of all the participants in each engagement along with status information; 2) the instrument panel of each F-15C cockpit which included the radar, radar warning receiver, and armament displays; and 3) the forward channel of out-the-window video for each F-15C cockpit. The plan view display, instrument panel displays, and radio communication were also recorded to video tape for mission debrief and further data analysis. In addition, the ECS included a data logger that recorded all the network communication protocols between simulators.

Ground threats, as well as additional threat and friendly aircraft, were provided by a computer- based automated threat engagement system (16). The ground threat portion of the automated threat engagement system (ATES) provided command and control functions (e.g., early warning radars and target assignment) and simulation of directed and autonomous surface-to-air missile batteries and anti-aircraft artillery with their radars. The aircraft portion of the ATES provided computer controlled air interceptors as well as formations of air-to-ground bombers. In addition, the ATES provided four computer controlled F-16s which were escorted by the manned F-15Cs during offensive counterair (OCA) sorties.

Scenario Design

The primary approach taken toward the measurement of SA was through scenario manipulation and observation of subsequent performance as recommended by Tenney, Adams, Pew, Huggins, and Rogers (17). Other approaches such as the use of explicit probes (9) were considered and finally rejected due to their lack of face validity for the study participants. Since we were using mission-ready F-15 crews, it seemed essential that we provide a simulation experience as realistic as possible. A week-long SA "evaluation" exercise was constructed that consisted of 9 sorties with 4 engagements per sortie. Sorties were arranged in a building block manner. Over the week, engagements increased in complexity in terms of numbers of adversaries, enemy tactics, lethality of ground threats, AWC support, etc.

A typical engagement scenario is presented in Figure 2. This depicts a defensive counterair (DCA) mission in which the objective of the two F-15s is to defend the home airfield. In this case, the attackers consist of two bombers accompanied by two fighters. The engagement begins at 80 nautical miles (nmi) separation in which the fighters are flying at 20,000 ft and the bombers at 10,000 ft. They are laterally separated by 10 nmi which makes them fairly easy to acquire on radar by the two F-15s. At 35 nmi, the fighters begin a corkscrew type of maneuver in which they rapidly descend to 3,500 ft. At this time, they will drop off of the F-15s radar screen. Upon completion of the maneuver, the fighters will trail the bombers as well as being at a much lower altitude. While the F-15s can easily continue tracking the bombers, it requires the crew to "predict" the actions of the fighters so that they may be quickly re-acquired on radar. At 15nmi, the bombers do a hard right turn and descend to 2,500 ft. At this time, the bombers will momentarily drop off the radar screen. Since the range is very close (10-12 nmi), it requires the crew to accurately "predict" the actions of the bombers and correctly use their radar so that they may be quickly re-acquired. The problem is further complicated in that the bombers and fighters will now "merge" in roughly the same airspace. If the fighters are ignored, then they can launch against the F-15s. If the F-15s "lock" their radar on the fighters, which will usually be the case at this point, then the bombers can continue toward the airfield "untargeted." Once the fighters are engaged, it is very difficult to re-acquire the bombers since they are low and will be flying away from the F-15s. If the F-15s fail to kill the fighters, the problem will only be compounded.

This example not only shows the approach taken toward the design of the mission scenarios, but also serves to illustrate our contention that SA is more than knowledge of the current situation. In operational environments, situation assessment and decision making are viewed as tightly coupled and are often difficult to separate. For the fighter pilot to be successful, he must not only be able to "build the big picture," but he must also translate his assessment into an employment decision. Often, the inability to make these critical employment decisions may lead to mission failure, despite a

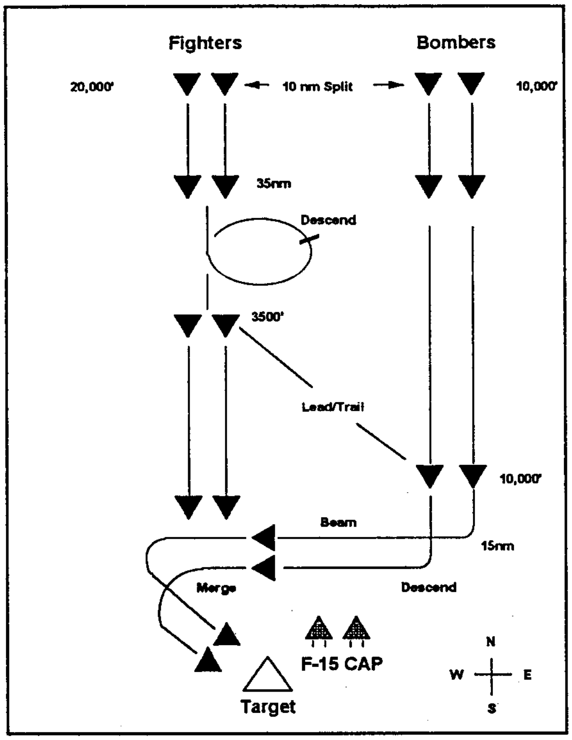


Figure 2. Typical Engagement Scenario

correct assessment of the situation. In the sample scenario, the key to success is to target and destroy the bombers prior to 15 nmi and then target the fighters. If the ranges become so close that all four threats must be dealt with simultaneously then the mission is likely to fail. It is through the careful design of such mission scenarios that the failure to incorrectly assess the situation or make incorrect employment decisions can be successfully inferred based upon the observation of pilot performance in the unfolding of the mission scenario.

Data Sources

Given the tremendous cost of gathering data on MR F-15 pilots, the approach was to gather as much as possible from a variety of sources. In our view, the most important data sources were the judgments and observations of two retired fighter pilots who possessed an in-depth understanding of the air combat domain. The same two SMEs were used throughout the year-long data collection effort. For each mission, the following procedure was followed. One of the SMEs would attend the mission briefing session conducted by the crew. During the conduct of each mission both SMEs observed mission performance. One of the SMEs also served as the mission director who was responsible for starting and stopping each engagement, communicating with the console operator, etc. During each engagement, each SME independently completed an observational checksheet to record pertinent events, notes, and outcomes. Upon completion of the four engagements comprising a single mission, one of the SMEs accompanied the crew to the debriefing room. The flight lead was responsible for conduct of the debriefing, although the SME was permitted to ask questions in an attempt to clarify the crew's understanding of the situation and purpose of their actions. Upon completion of the debrief, the two SMEs discussed each engagement, and completed a consensus performance rating scale consisting of the 24 behavioral indicators of SA related to F-15 mission performance. The SMEs also produced a written critical events analysis for each mission which attempted to identify those events that, in their opinion, affected the outcome of the mission and were indicative of the crew's SA.

A variety of other data were also gathered. These included mission events and outcomes such as weapons firings, kills, etc. Using the data logger in the ECS, the digital data passed over the network was recorded, whereby each engagement could be reconstructed. The videos recorded and used for debriefing were also archived. Additionally, eye movement data were recorded for the four engagements flown on the last mission. And finally, all participants were also asked to "critique" the simulation and also give opinions regarding its potential for training.

Results

The data gathered in this study provided information appropriate for the first two stages of the evaluation model presented earlier.

Utility Evaluation

Two types of user opinion data were gathered. First, pilots rated the training benefit for various pilot experience levels. And second, pilots completed an open-ended questionnaire pertaining to the overall value of the simulation and how it might best be used.

The results of the ratings of potential training benefits are provided in Figure 3. These data clearly indicate that positive opinions were expressed by the study participants on the value of this type of simulation for training. The potential training was considered beneficial for all levels of qualification. It is of interest to note that training was considered highly beneficial for four-ship flight leads, despite the fact that the MULTIRAD simulation facility provided training for only a flight lead and wingman. As expected, higher benefit ratings were given to pilots upgrading into a given qualification level.

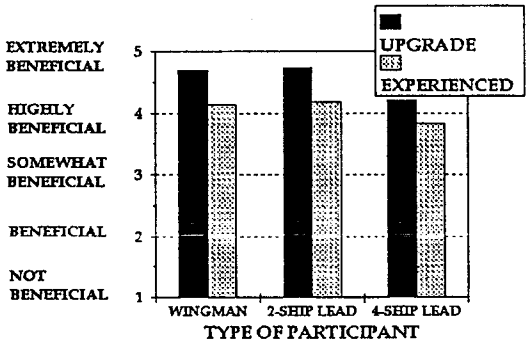


Figure 3. Rated Benefit of Training for Various Levels of Experience

Opinions expressed in the open-ended questionnaire were also quite positive. Although qualitative, they provide additional insight into the potential focus of training using multiship simulation and how it might be employed. In particular, mention was made of using such training as a means of enhancing both situation assessment and decision-making skills. It was also frequently noted that there was tremendous value in learning flight leadership and resource management skills. In terms of the location of such simulation, the overwhelming consensus was that they would be of most value within the operational units. This was not too surprising since each unit now has the operational version of the cockpits used in the present investigation. However, they are stand-alone and non-visual, and as such their training capability is fairly limited. In contrast, the networking of such devices within a realistic combat environment increases the potential greatly. The bottom line from the utility data is that the participants considered multiship simulation as a tool with high training potential.

Performance Improvement

While positive user opinion is a necessary prerequisite for effective training, in itself, it is insufficient validation (5). At the next stage of the evaluation model, it is necessary to demonstrate improved performance within the simulation environment as a function of practice. In other words, it is necessary to show that learning has occurred. It should be pointed out that it was never the intent, at the outset of the study, to demonstrate performance improvements. It must be emphasized that the sole purpose was to develop a set of simulation scenarios that could be used to assess SA within a combat environment. As such, normal training interventions were not permitted. For example, during the debrief, pilots were permitted to only view their own in-cockpit displays and not the planned view display. Moreover, the two SMEs were not permitted to provide any type of feedback to the pilots regarding their performance.

However, data from the ninth mission did permit some comparison since identical scenarios had been flown earlier in the week. The ninth mission was designated the "eye track" mission in which eye movement data were recorded. For these scenarios, an eye tracker computed point of gaze and was displayed against the background scene as determined from a scene camera mounted on the pilot's helmet. The resulting video signal replaced the second cockpit display within the ECS. This permitted the crews to debrief the final mission using three integrated displays, the planned view of the fight, their own cockpit display, and the eye-tracked display which portrayed point of gaze against the background scene. Although not central to this paper, it should be mentioned that very positive opinions were expressed by the pilots regarding the potential of eye movement recordings as a feedback tool for training. It was viewed as potentially useful for the earlier stages of training and, in particular, for the diagnosis of problems of students encountering difficulty. It could potentially provide a solution to the continuing problem of training for single-seat aircraft in which instructors complain that diagnosis is difficult when one cannot see where the student is looking.

Two scenarios, a 2 V 2 DCA mission and a 2 V 4 OCA mission, were flown during the middle of the week and then again on the last mission. A comparison of performance is presented in Figure 4. In both cases, performance on the last mission was improved. However, only the 2 V 2 DCA mission was found to be statistically significant.

It should be recalled that the scenarios were designed to increase in difficulty over the week. Consequently, if one simply plots the Observer SA scores across missions, there is generally a downward trend. To obtain an estimate of what the

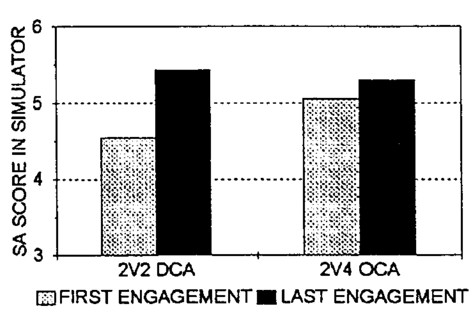


Figure 4. Effects of Practice on Observer SA Ratings

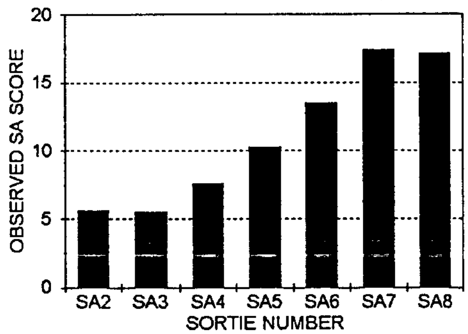


Figure 5. SA Scores Weighted for Scenario Difficulty Across Missions

curve might look like assuming "equal difficulty" of all scenarios, a magnitude estimation procedure was undertaken to scale the difficulty of the scenarios. Raters included the two SMEs and another in-house F-15 pilot who had occasionally served as wingman in the course of the study. Only missions 2 through 8 were included since mission 1 was a "familiarization" sortie and mission 9 was the eye track sortie. These difficulty weightings were then applied to the mean observed SA scores for each mission. The results are presented in Figure 5.

It is clear that when the scores are weighted for scenario difficulty, the resulting curve suggests that performance improved over the week. Again, it should be cautioned that the procedures followed were not the most appropriate for a conduct of a rigorous test of learning within the simulation environment. However, when such data are coupled with the very strong pilot opinions that they had received valuable training, it seems reasonably safe to conclude that learning had occurred over the week.

Discussion

The results obtained in the present study confirm previous efforts attempting to measure the training benefits of interactive air combat simulation. From a user's perspective, the data are very clear regarding the potential value of such simulation for training. The 63 MR F-15 pilots overwhelmingly considered such training to be of value. Although such anecdotal evidence is often considered suspect from a scientific perspective, it is nevertheless an absolute prerequisite for effective training. Unless there is user acceptance, the resulting training will be of marginal value regardless of the device's inherent potential.

In addition to the opinion data, there is evidence that performance did improve within the simulation environment; in other words, learning did occur. Again, it should be pointed out that the amount of improvement was probably "minimized" due to the evaluative orientation of the investigation. When identical scenarios were flown early and late during the week, the performance on the second repetition was better. Additionally, when scenario difficulty is assumed constant, the resulting weighted scores show improvements. These data combined with the fact that the study participants expressed opinions to the effect that their proficiency had improved leave little doubt that learning had occurred.

Although the data clearly indicate that the end user expresses very positive opinions toward the value of multiship simulation and that learning occurs, there still remains the issue of transfer of training. Does such training transfer to other simulation environments (Stage 3 of the Evaluation Model) and does it transfer to the real world (Stage 4 of the Evaluation Model) which represents the "acid test?" Clearly, the data gathered in this study do not bear upon these issues. Future research investigations are now being planned to answer these questions.

At present, the question of training benefits of interactive air combat simulation is largely answered by one's personal view of simulation. For the "believer," evidence to date is strong enough to warrant the conclusion that training will be effective. In fact, given the previous transfer of training research that has already been conducted (18,5) there is little reason to suspect that such training within a multiship simulation environment would not have a positive effect upon subsequent performance in the air. Yet, for the "skeptic," no definitive evidence has been presented.

References

1. United States Air Force Scientific Advisory Board. (1992, December). Report of the Support Panel of the Air Force Scientific Advisory Board Ad Hoc Committee on Concepts and Technologies to Support Global Reach - Global Power 1995 - 2020. Washington, DC: Author.

2. Gray, T.H., Edwards, B.J., & Andrews, D.H. (1993). A survey of F-16 squadron level training in PACAF (AL-TR-1993-0041). Williams Air Force Base, AZ: Armstrong Laboratory, Aircrew Training Research Division.

3. Houck, M.R., Thomas, G.S., & Bell, H.H. (1991). Training evaluation of the F-15 advanced air combat simulation (AL-TP-1991-0047). Williams Air Force Base, AZ: Armstrong Laboratory, Aircrew Training Research Division.

4. Alluisi, E.A. (1991). The development of technology for collective training: SIMNET, a case history. Human Factors, 33, 343-362.

5. Bell, H.H. & Waag, W.L. (1994). Evaluating the Effectiveness of Flight Simulators for Training Combat Skills: A Review. Submitted to International Journal of Aviation Psychology.

6. Deitchman, S.J. (1988). Preliminary exploration of the use of a warfare simulation model to examine the military value of training (IDA Paper P-2094). Alexandria, VA: Institute for Defense Analysis.

7. Thomas, G.S., Houck, M.R. & Bell, H.H. (1990). Training evaluation of air combat simulation (AFHRL-TR-90-30). Williams AFB AZ: Air Force Human Resources Laboratory, Operations Training Division.

8. Carroll, L.A. (1992). Desperately seeking SA. TAC Attack (TAC SP 127-1) 32: 5-6.

9. Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. In Proceedings of the Human Factors Society 32nd Annual Meeting (pp. 97-101). Santa Monica, CA: Human Factors Society.

10. Fracker, M.L. (1988). A theory of situation assessment: implications for measuring situation awareness. In Proceedings of the Human Factors Society 32nd Annual Meeting (pp. 102-106). Santa Monica, CA: Human Factors Society.

11. Houck, M.R., Whittaker, L.A., & Kendall, R.R. (1993). An information processing classification of beyond-visual-range intercepts. (AL/HR-TR-1993-0061). Williams AFB AZ: Armstrong Laboratory, Aircrew Training Research Division.

12. Waag, W.L. & Houck, M.R. (1994). Tools for assessing situational awareness in an operational fighter environment. Aviation Space and Environmental Medicine. 65 (5, Suppl.):A13-A19.

13. Carretta, T.R., Perry, D.C., & Ree, M.J. (1994). Prediction of situational awareness in F-15 pilots. Submitted to International Journal of Aviation Psychology.

14. Gehl, T.L., Rogers, R.L., Miller, M.A., & Rakolta, J. (1993). Interdependence of training utility and network performance using the Armstrong Laboratory Multiship Research and Development System. In, Proceedings of 15th Industry/Interservice Training Systems Conference, Orlando, FL: National Security Industrial Association.

15. Platt, P. & Crane, P. (1993). Development, test, and evaluation of a multiship simulation system for air combat training. In, Proceedings of 15th Industry/Interservice Training Systems Conference, Orlando, FL: National Security Industrial Association.

16. Rogers, B. (1992). Tactical air threat system for a distributed simulation network. In, Proceedings of 14th Industry/Interservice Training Systems Conference, San Antonio, TX: National Security Industrial Association.

17. Tenney, Y.J., Adams, J.J., Pew, R.W., Huggins, A.W.F, & Rogers, W.H. (1992). A principled approach to the measurement of situation awareness in commercial aviation (NASA Contractor Report 4451). Langley VA: National Aeronautics and Space Administration.

18. Waag, W.L. (1981). Training effectiveness of visual and motion simulation. (AFHRL-TR-79-72). Williams AFB AZ: Air Force Human Resources Laboratory, Operations Training Division.